

Reference Magnetism and Macro-Naturalism

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation *with research distinction in Philosophy*
in the undergraduate colleges of The Ohio State University

by
Gabbrielle M Johnson

The Ohio State University
June 2011

Project Advisor: Ben Caplan, Department of Philosophy

Reference Magnetism and Macro-Naturalism¹

1. Introduction

Realist theories of reference address the question of what objective feature it is in virtue of which a word, e.g., ‘Columbus’, refers to the object that it does, i.e., Columbus, rather than, say, Portland. Hilary Putnam (1977: 123-140) argues that such features don’t exist, so there is no “semantic glue” that ties our words to their referents. The purpose of this paper is to trace the attempts made by philosophers such as David Lewis and Theodore Sider to find this glue. By exploring the dispute between these two philosophers and their opponents, I hope to illustrate the basic requirements on a theory attempting to secure reference and also illustrate why their theories don’t satisfy those requirements.

2. Putnam

2.1 Charity

I will start, as Putnam does, with the notion of *intentionality*: the idea that words and sentences in a given language are, in fact, about the world; they have intentional content. The notion of intentionality is important for any theory of reference because it motivates the assumption that words not only refer, but also they refer to objects in the world. That is, for any particular language, say L, L is informative and about the world. To interpret L is to determine what particular objects are assigned to words in L. The interpretation that is intuitively accurate, i.e., one that assigns the right objects to names and the right extensions to predicates, is the *intended interpretation*. A *model* of a theory is an interpretation of the language in which the sentences of the theory are written and that makes those sentences true. Various theories of reference differ about what principles guide reference assignments and, thus, in virtue of what words in L refer to the objects that they do, i.e., the objects assigned under the intended interpretation.

¹ I would like to thank Ben Caplan for an immense amount of support and insight during the process of completing this thesis.

For the purposes of my project, we can assume that the intended interpretation is one that is a model of a theory that contains sentences in L that express the beliefs of speakers of L. We do this because we assume that, *ceteris paribus*, speakers of L are rational and, therefore, that, *ceteris paribus*, they have true beliefs. Thus, we have the intuition that, *ceteris paribus*, utterances in L that express those beliefs are true.²

Granting the assumption that speakers intend to refer and use predicates in such a way that those sentences that express their beliefs are true, we can characterize an initial principle for reference—call this ‘the Principle of Charity’ or ‘PC’:

The Principle of Charity (PC): The intended interpretation of a language L is one on which referents are assigned to expression in L in a manner that makes the sentences of a theory T true, where T is the set of sentences of L that express the beliefs of speakers of L.³ In the next part of this paper, I will provide a counterexample that will motivate revision to this principle.

2.2 Counterexample One: Approximation World

Imagine a scenario in which Mario is asked to partake in a guessing contest. The task of the contest is to guess the approximate number of Jellybeans in ‘The Mason Jar.’⁴ The only way to win the contest is to be the individual whose guess comes closest to the precise number of beans without going over that number. Mario, having an esteemed reputation for confectionary speculations, confidently asserts

(1) ‘There are at least 742 Jellybeans in ‘The Mason Jar.’

² My interest for this paper is to investigate the various theories that attempt to offer principles by which we can determine reference. A speaker’s beliefs presumably have intentional content and, therefore, one might ask in virtue of what those beliefs latch on to the objects that they do. However, for my purposes, it is sufficient to assume that a speaker’s beliefs adequately correspond to the sentences in L that are true.

³ Philosophers who invoke the notion of charity disagree about how best to characterize the principle of charity. For reasons discussed above, my characterization relies heavily on the assumption that speakers of L have true beliefs; however, we will return to the details of this discussion in §4.2.

⁴ For simplicity, throughout my paper, ‘The Mason Jar’ will act as a proper name.

As it turns out, there are actually 927 Jellybeans in the jar that is in front of Mario at the time he utters (1), 742 of which were hand-picked earlier by David Klein at the Jellybelly headquarters in Fairfield, California.

If we interpret Mario's utterance in an intuitively correct manner, we would choose an interpretation that assigns to 'The Mason Jar' the jar that is in front of Mario at the time that he utters (1) and that assigns to 'Jellybeans' the set of all Jellybeans⁵—call this the 'I-interpretation'.⁶ On the I-interpretation, Mario has not guessed the precise number of Jellybeans; however, luckily for him, he has not guessed over 927 and no other contestant guessed a number closer to 927. Therefore, Mario has won the contest.

There are, of course, other ways of interpreting Mario's utterance. For example, we might adopt an interpretation that assigns to 'The Mason Jar' the jar that is in front of Mario at the time that he utters (1) and that assigns to the predicate 'Jellybean' the set of Jellybeans that have the property *being at some point hand-picked by David Klein at the Jellybelly headquarters in Fairfield, California*—call this the 'A-interpretation'.⁷ On this interpretation, Mario also manages to live up to his reputation.

This example, in effect, illustrates Putnam's model-theoretic argument. That is, both the I-interpretation and the A-interpretation score equally well with regard to PC. Moreover, there are countless interpretations available that will satisfy PC, thereby yielding *some* (though not necessarily intuitive) intentional content. The intended interpretations that seem at odds with our intuitions are traditionally referred to as "bizarre" or "crazy."⁸

⁵ Following Lewis (1986), I take *all* Jellybeans to include the Jellybeans located in all possible worlds. This becomes important in §3.2.

⁶ 'I' is for 'intuitive'.

⁷ 'A' is for 'approximation'.

⁸ Williams (2007: 377), Hawthorne (2007: 1), and Sider (forthcoming: 29).

One way of illustrating the deficiencies of bizarre interpretations is testing whether they can account for counterfactual robustness, i.e., bizarre interpretations fail to account for counterfactual evidence.⁹ For example, if another contestant were to have guessed a number between 742 and 927, then Mario would have lost the contest. Intuitively, this counterfactual statement is true. Moreover, on the I-interpretation of these utterances, it is true.¹⁰ However, on the A-interpretation of the contestants' utterances, the statement is false. Therefore, it seems that the A-interpretation is deficient, but with PC alone an interpreter has no way of motivating the I-interpretation over the A-interpretation.

Therefore, to block bizarre interpretations such as the A-interpretation, we'll need to revise PC, to give us a different way of ranking some interpretations as more suitable than others. In the following section, I will discuss Lewis's attempt to do this by positing his *eligibility constraint*.

3. Lewis

3.1 Eligibility Based on the Naturalness of the Reference Relata

To block bizarre and unintended interpretations, e.g., the A-interpretation, Lewis proposes an additional constraint called 'eligibility'. Lewis (1983a: 371) says

I ... propose that the saving constraint concerns the referent – not the referrer and not the causal channels between the two Reference consists in part of what we do in language or thought when we refer, but in part it consists in eligibility of the referent. And this eligibility to be referred to is a matter of *natural* properties. (Italics added)

In other words, what Lewis calls 'natural' are those properties that are more eligible candidates for reference and, thus, act as magnets for reference.

⁹ Sider (forthcoming: 36-37) uses a similar strategy for illustrating the weakness of charity.

¹⁰ According to Putnam (1977), statements such as this counterfactual, PC, and other statements intended to guide theories of reference are just other sentences in T. Lewis (1984) responds to this objection; cf. §3.1 and n. 16. For the sake of brevity, assume that the interpretation of this counterfactual statement, other counterfactuals to follow, and the various principles of reference are not considered suspect.

These natural properties, which Lewis (1983a: 364) considers to be an “elite minority of special properties,” are useful only insofar as we can distinguish them from the majority of properties, which are “gerrymandered and miscellaneous ... [and] superfluous in characterizing the world.” To help with this distinction, Lewis contrasts properties that he calls *sparse* and those that he calls *abundant*. The abundant properties are those that are “gruesomely gerrymandered,” “miscellaneously disjunctive,” and “carve things up every which way.”¹¹ Conversely, of the sparse properties he says

Sharing of them makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are *ipso facto* not entirely miscellaneous, and there are only just enough of them to characterize things completely and without redundancy.¹²

These sparse properties, then, constitute the natural properties.¹³ Therefore, following Lewis (1983a), the property *being an electron* might be natural, but the property *detecting the Higgs boson* is not. However, Lewis (1986) suggests that naturalness might come in degrees, making some properties more natural than others. In other words,

Probably it would be best to say that the distinction between natural properties and others admits of degree. Some few properties are *perfectly* natural. Others, even though they may be somewhat disjunctive or extrinsic, are at least somewhat natural in a derivative way, to the extent that they can be reached by not-too-complicated chains of definability from the perfectly natural properties.¹⁴

¹¹ Lewis 1986: 59.

¹² Lewis 1986: 60; italics in original.

¹³ It might be helpful to distinguish between *fundamental*, *microphysical*, and *natural* properties. The *natural* properties are those that carve at the joints. The *microphysical* properties are those posited by physics. The *fundamental* properties are those that obtain, but not in virtue of other properties obtaining. Lewis (2009) refers to the fundamental properties as ‘perfectly natural’ and Sider (forthcoming) equates fundamental properties with those envisaged by physics, i.e., microphysical properties. However, Schaffer (2004) argues that natural properties can occur at many levels of nature. For more on this discussion, see Lewis 2009: 204-205; Sider forthcoming: 119-123, 141-143; and Schaffer 92-93, 97-98. Also, Lewis (1986, 2009) takes naturalness to come in degrees; whereas, Sider (forthcoming) takes naturalness to be *absolute* (to be discussed below).

¹⁴ Lewis 1986: 61; italics in original.

Following this shift, *detecting the Higgs boson* is not perfectly natural, yet it is still more natural than the property *being a 17-mile-long particle accelerator near Geneva, Switzerland that might have detected the Higgs boson*.

Here, Lewis has established a certain set of privileged properties that exist independently of language¹⁵ and are most eligible to be the referents of our words. We can now revise our initial principle of reference to echo this eligibility constraint—call this principle ‘Charity Plus Lewis Eligibility’ or ‘CPLE’:

Charity Plus Lewis Eligibility (CPLE): The intended interpretation of a language L is one on which (i) referents are assigned to expressions in L in a manner that makes the sentences of a theory T true, where T is the set of sentences of L that express the beliefs of speakers of L and (ii) the referents thus assigned are eligible, where eligibility is determined by naturalness, and perfectly natural properties are both microphysical and fundamental. By appealing to CPLE, we are able to block the A-interpretation presented by the Approximation World, because presumably the property *being a Jellybean* is more natural than the property *being a Jellybean that was at some point hand-picked by David Klein at the Jellybelly headquarters in Fairfield, California*. In other words, the I-interpretation, which assigns the set of Jellybeans to the expression ‘Jellybean’, maximizes charity—satisfying (i)¹⁶—while preserving objective naturalness—satisfying (ii). Conversely, on the A-interpretation, we satisfy (i) but only by sacrificing (ii). Therefore, according to CPLE, the I-interpretation is the intended interpretation.

It is not entirely clear, however, that CPLE can adequately block all bizarre interpretations. In the next section, I will again give a counterexample that will motivate further revision to this principle of reference.

¹⁵ This is partially a response to Putnam (1977), who objects that any attempt to offer an additional constraint must be done in some language that would itself be subject to the same worry about finding the semantic glue as the theory it attempts to save. Lewis (1984) suggests, instead, that if the additional constraint is independent of language, it’s presumably exempt from such arguments; cf. n. 9.

¹⁶ As demonstrated in §2.2.

3.2 Counterexample Two: Belief World

In “Craziness and Metasemantics,” John Hawthorne (2007) presents several scenarios with the intention of undermining Lewis’s eligibility constraint. The most pernicious of these is an example in which there is a twin world—call this ‘Twin Earth’—where events correspond exactly with the beliefs of speakers on Earth.¹⁷

For example, again imagine that Mario is asked to guess the number of Jellybeans in The Mason Jar. This time, however, also imagine that there exists a distant galaxy with a Twin Earth that corresponds to Earth in all respects except for one: rather than there being 927 Jellybeans in the twin-jar, there are exactly 742 Jellybeans. Again, there are a variety of interpretations available for Mario’s utterance. For example, we might say that when Mario says ‘The Mason Jar’ he is referring to the twin-jar—call this the ‘T-interpretation’.¹⁸ The I-interpretation and the T-interpretation score equally well with regard to charity and, as Hawthorne (2007: 428) points out, “[T-interpretation] comes out no worse on the score of eligibility... . So neither charity nor eligibility can explain the craziness of the interpretation.”

Presumably Mario’s Twin, the twin-jar, and the twin-beans all have as many natural properties as the earthly Mario, jar, and beans do. Therefore, according to CPLE, the T-interpretation is no worse than the I-interpretation. However, intuitively the T-interpretation is bizarre and, thus, incorrect.

4. Sider

4.1 Eligibility Based on the Naturalness of the Reference Relation

¹⁷ For Hawthorne, the Twin Earth counterexample has an interpretation that scores higher with regard to charity because the events of Twin Earth correspond more closely with the beliefs of speakers on Earth than Earth events do; however, it is sufficient for my purposes to use an example in which the Twin Earth interpretation ties with the I-interpretation in terms of charity.

¹⁸ ‘T’ is for ‘Twin Earth’.

Even though Lewis's revised principle of reference has not been able to block bizarre interpretations, e.g., interpretations on which our referents are located galaxies away, Theodore Sider remains sympathetic to Lewis's project. Moreover, he suggests that by revising our notion of eligibility we can salvage it.

The problem, according to Sider, is that an adequate theory of reference will focus, not on the eligibility of the referent, but rather on the eligibility of the relation between the referrer and the referent.¹⁹ According to Sider (forthcoming: 33),

If reference-based explanations are to be genuinely explanatory, then the reference relation must be a joint-carving one. ... And this excludes the bizarre interpretations. For only a wildly non-joint-carving relation would relate a linguistic population to the semantic values of a bizarre interpretation.

This allows us to block bizarre interpretations such as the T-interpretation because a relation between Mario and the jar on Twin Earth is not suitably natural.²⁰ We are now in a position to revise our revised principle of reference—call this principle 'Charity Plus Sider Eligibility' or 'CPSE':

Charity Plus Sider Eligibility (CPSE): The intended interpretation of a language L is one on which (i) referents are assigned to expressions in L in a manner that makes the sentences of a theory T true, where T is the set of sentences of L that express the beliefs of speakers of L, and (ii) the relation between those speakers and the referents thus assigned is eligible, where eligibility is determined by naturalness, and natural properties are both microphysical and fundamental.

I will now present three counterexamples that are designed to put pressure on this revised principle of reference.

4.2 Counterexample Three: Stranger than Fiction World²¹

¹⁹ I will return later to the question of whether this is in keeping with other tenets of Sider's work, and here I am not particularly concerned with the deeper motivations behind such a shift; cf. §5.1.

²⁰ Under Lewis's theory, a model should include in the extension of 'Jellybean' not only worldly Jellybeans, but also otherworldly or twin jellybeans. This is perhaps why Lewis did not pursue eligibility as a matter of natural relations.

²¹ This title is borrowed from the movie *Stranger Than Fiction*, where a similar scenario occurs.

To demonstrate that CPSE can be flawed in the same manner as PC and CPLE, I again invoke the story of Mario and the Jellybeans. However, before continuing, I must admit that I borrowed the name ‘Mario’ from a friend. And, unbeknownst to me, it might be the case that, at the very time that I am telling the story of the guessing contest, Mario is actually being asked to guess the precise number of Jellybeans in a mason jar. Mario asserts (1), the jar actually contains 927 Jellybeans, but he wins just the same. Here, a story that I have taken to be fiction really corresponds to some actual events. In this case, does my word ‘Mario’ refer to the actual person despite my having no intention to do so? Call this possible interpretation that assigns the actual individual Mario to my word ‘Mario’ the ‘F-interpretation’.²²

A case similar to this is discussed by both David Lewis (1978: 39-40) and Saul Kripke (1980: 157-159) and involves Arthur Conan Doyle’s stories about Sherlock Holmes. Again we can imagine that, while Conan Doyle was writing what he took to be fiction, there actually existed a man by the name of ‘Sherlock Holmes’ who lived in London and solved mysteries. Therefore, when Conan Doyle utters (2), the real Sherlock Holmes actually exercised before eating breakfast.

(2) Sherlock Holmes exercised before he ate breakfast.²³

As Lewis (1978: 39) remarks, such a case is “improbable, incredible, but surely possible!”

Unlike in previous examples where there have been proximal objects that stand in close relations to the speaker, e.g., the Jellybeans and jar that are spatiotemporally nearer to Mario than the Twin objects are, in both instances here there are no better candidates for being the referent of ‘Sherlock Holmes’ and ‘Mario’ than the real-life individuals who go by these names. Similarly, there is no better candidate for being the relation between speakers and the referents of those words than whatever causal or explanatory relation exists between me and the real-life Mario or between Conan Doyle and the real-life Holmes. Therefore, according to CPSE, F-interpretation is the intended

²² ‘F’ is for ‘fiction’.

²³ See Conan Doyle 1904.

interpretation. Despite this outcome, however, we still take the F-interpretation to be highly counterintuitive. For example, Lewis (1978: 39) says

Let us assume that Conan Doyle indeed wrote the stories as pure fiction. He just made them up. He had no knowledge of anyone who did the deeds he ascribed to Holmes, nor had he even picked up any garbled information originating in any such person Now consider the name “Sherlock Holmes,” *as used in the stories*. Does the name, so used, refer to the man whom Conan Doyle never heard of? Surely not! (Italics in original)

Although Lewis here is referring to names used in the fiction, e.g., Watson’s use of ‘Sherlock Holmes’, the point can be extended to real-life uses of the name. Kripke (1980: 157-158), for example, straightforwardly rejects the claim that sentences outside of the fiction that use names from within the fiction could coincidentally be about actual people:

The mere discovery that there was indeed a detective with exploits like those of Sherlock Holmes would not show that Conan Doyle was writing *about* this man; it is theoretically possible, though in practice fantastically unlikely, that Doyle was writing pure fiction with only a coincidental resemblance to the actual man. (See the characteristic disclaimer: ‘The characters in this work are fictional, and any resemblance to anyone, living or dead, is purely coincidental.’) (Italics in original)

However, to evaluate the extent to which the Stranger Than Fiction World can serve as a counterexample, we must return to the details of charity that we initially glossed over. Again, according to PC and, consequently, CPLE and CPSE, referents are assigned to expressions in L in a manner that makes the sentences in T true—so the question is whether my sentences about someone I call ‘Mario’ express my beliefs and, thus, belong in T. There is a sense in which we would want a sentence such as (2) to come out true while sentences such as (2*) come out false.

(2*) Sherlock Holmes exercised on the treadmill before he ate breakfast.²⁴

This is because, given the fiction, we do not believe Sherlock Holmes exercised on a treadmill.

Similarly, we might attempt to restrict the assertions that I can make about my fiction so that sentences such as ‘David Klein invented Jellybelly Jellybeans’ come out true, while others such as

²⁴ The treadmill was not used for aerobic exercise until around 1950, approximately 50 years after the fiction was written.

‘Calvin Klein invented Jellybelly Jellybeans’ come out false. It seems plausible, then, that how we talk about fiction is determined, in part, by the relevant facts presented within the fiction, which shape our beliefs. Therefore, when I make assertions such as “Mario has won the contest,” that sentence belongs in T; it expresses one of my beliefs. However, this coupled with CPSE leads to the unfortunate conclusion that my word ‘Mario’ refers to the real-life individual. But, in the event that we would not want such sentences to reflect our beliefs independent of the fiction, we can move on to the next counterexample.

4.3 Counterexample Four: Hallucination World

Charity requires the interpretation of L to render all sentences in L that express the beliefs of speakers true, but it’s unclear whether we should take sentences about fiction to be expressing beliefs. Therefore, in this sub-section I hope to present a case in which it is clear that the utterance expresses one of the speaker’s beliefs and that the bizarre interpretation scores better according to CPSE.

Imagine a scenario in which Mario is utterly fed up with guessing any amount of Jellybeans in any jars, either those spatially near him or those existing in other worlds. He is so unnerved by the repeated requests to demonstrate his speculative prowess that the mere sight of anything Jellybean- or Mason-jar-related jolts him into a state of hysteria, invoking rather odd hallucinations.

Despite his nonattendance, the annual guessing contests continue without Mario. However, one fateful day Mario has the extreme misfortune of unknowingly walking in on such a contest taking place. Upon seeing the jar of Jellybeans, Mario is sent into a fit of distress in which he hallucinates a past experience of guessing the Jellybeans years before and hysterically asserts (3).²⁵

(3) There are at least 927 Jellybeans in The Mason Jar.

²⁵ I take it that Mario’s intention in uttering (3) is to win the contest that he’s hallucinating, which happens to be reminiscent of a past contest, and not to win the contest he has happened upon accidentally.

And, coincidentally enough, the jar that is in front of Mario at the time that he utters (3) does in fact have 927 beans in it.²⁶ Therefore, we could create an interpretation on which the expression ‘The Mason Jar’ refers to the jar that is in front of Mario at the time that he utters (3)—call this ‘H-interpretation’.²⁷

Just as in the Stranger than Fiction World, where the most viable referents and relations were the real-life persons and the relations existing between them and the referrers, in the case of Mario’s hallucination, the real-life perceptual and causal relations between Mario and the jar are most viable. Therefore the H-interpretation is the intended interpretation according to CPSE. However, intuitively, this is a bizarre interpretation. Consider that if this hallucination had taken place in Mario’s bedroom, then it’s clear Mario would have lost the contest. Moreover, the hallucination is the same whether he is in his bedroom or in the contest hall. Therefore, Mario’s location does little to affect the content of the hallucination and the mere coincidence that Mario’s hallucination *corresponds* to the real world does little to affect our intuitions.

According to CPSE, however, the H-interpretation is the intended interpretation, since (a) Mario correctly guessed the number of beans in the jar that is in front of him at the time that he utters (3), (b) (3) expresses his beliefs (he’s hallucinating and doesn’t know it, so he believes that what he hallucinates is real), and, moreover, (c) the jar that is in front of Mario at the time that he utters (3) stands in natural relations—both causal and perceptual—to Mario.²⁸

4.4 Counterexample Five: Star World

Imagine a scenario in which Mario, who is located on Earth, utters the following:

(4) Betelgeuse is the second brightest star in the constellation Orion.

²⁶ Perhaps the repetition in the amount of beans explains how it is that Mario came to have the reputation that he does.

²⁷ ‘H’ is for ‘hallucination’.

²⁸ Granted, this is a *deviant* causal chain. But there is definitely some causal relation between Mario and the jar that is in front of him at the time that he utters (3), which is all that is needed for now to satisfy CPSE. More on this discussion is located in §5.1 below.

We can adopt an interpretation that assigns to the expression ‘Betelgeuse’ the star Betelgeuse. However, current estimates place the location of this star to be at least 1,300 light-years away, allowing that Betelgeuse has already exploded.

For the purposes of this paper, let us assume that Betelgeuse has, in fact, gone supernova and no longer exists. However, the light traveling away from this point in the galaxy is moving at such a speed that Mario is currently seeing the light from 1,300 years ago, when the star still existed—call this a ‘perceptual lag’. Certainly Mario is not attempting to tell a fictional story when he utters (4); rather, we can reasonably take him to be uttering a sentence that expresses one of his beliefs. This leaves us with two possible interpretations: we could interpret Mario’s utterance in such a way that his word ‘Betelgeuse’ refers to (A) the object that existed a millennium ago and was located 1,300 light-years away or (B) the light that is currently stimulating his retina.

Since Betelgeuse is available to Mario only in virtue of his *seeing* it, i.e., his interacting with the photons traveling from the star, it seems that Mario has a more direct and, thus, more natural relation with the light that is currently stimulating his retina than he has with the object that existed a millennium ago and was located 1,300 light-years away. Whereas, on the I-interpretation ‘Betelgeuse’ refers to (A), we can call the interpretation that assigns (B) to ‘Betelgeuse’ ‘S-interpretation’.²⁹ Moreover, we could assign to the predicate ‘star’ in S-interpretation the property *being a particle of light traveling away from the object that existed a millennium ago and was located 1,300 light-years away*,³⁰ thus making (4) true. According to CPSE, the S-interpretation is the intended interpretation because it is just as charitable as the I-interpretation and the relation between Mario and (B) is more natural than the relation between Mario and (A).

²⁹ ‘S’ is for ‘star’.

³⁰ This is a very rough characterization of what such an assignment could be and we would need to reinterpret other names such as ‘constellation’; however, for this example to work, it is enough that there is *some* assignment that will satisfy (4) being true on the S-interpretation, which, according to Putnam (1977), presumably there is.

It seems, however, that the S-interpretation is highly unintuitive. For example, consider that on S-interpretation if Mario had been blind at the time that he uttered (1), the sentence would have been false. This is because there is no light that is stimulating his retina (even though there are objects with the property *being a particle of light traveling away from the object that existed a millennium ago and was located 1,300 light-years away*).³¹ However, we have the intuition that Betelgeuse is the brightest star in the Orion constellation regardless of whether Mario can see it.

Moreover, it seems that, although the perceptual lag is more noticeable on the grand scale, it must still exist to a smaller degree for every object reflecting light. Therefore, adopting the S-interpretation has dangerous consequences for a variety of utterances within Mario's language, not just those concerning stars that are light-years away. For example, according to the S-interpretation, when Mario utters (1) his expression 'The Mason Jar' actually refers to the light that has traveled from the jar that was in front of Mario a nanosecond before he uttered (1).³² This is because the light that acts as the mediator between Mario and the jar stands in a more natural relation—causal or perceptual—to Mario than the jar does to Mario.³³ Again, CPSE does not provide the necessary means to block this obviously bizarre interpretation.

5. Hawthorne and Schaffer

5.1 Eligibility Based on Macro-Naturalism

The theories presented thus far by Lewis and Sider have built on the original notion of charity with the additional constraint of eligibility. The notion of eligibility, however, is different for each. To better see the contrast between the two views, we must distinguish between the question of (a) whether it is the eligibility of the referent or the relation that matters for reference and the question of (b) whether eligibility is a relative or an absolute notion. For question (a) Lewis claims that the

³¹ Mario is referring to the light that is stimulating his retina, as mentioned in (B).

³² Assuming Mario is standing one foot away from the jar.

³³ Moreover, every form of sense perception has some complex route of stimulus that travels to the brain. Therefore, every object a speaker perceives there is some perceptual intermediary that stands in a more natural and eligible relation to her. Thus, every object that a speaker perceives stands in a less natural and hence less eligible relation to her than does some perceptual intermediary.

naturalness of the referents is what determines eligibility; whereas Sider claims that the naturalness of the reference relation is what determines eligibility. In regard to question (b), Lewis argues that naturalness is relative and suggests that we can grade naturalness by the definitional distance of a predicate (in a canonical language) that expresses one property and a predicate (in that language) that expresses a perfectly natural property.³⁴ On the other hand, Sider claims that naturalness is an absolute notion; therefore, things are either perfectly natural or not natural. Although they disagree about (a) and (b), both theories adopt a notion of naturalness according to which the perfectly natural properties are those that exist on the microphysical ground floor, call this ‘Micro-Naturalism’.

Micro-Naturalism: Natural entities are microphysical.

Sider (forthcoming: 119-123) also posits the following principle called ‘purity’:

Purity: Natural notions can be applied only to other natural notions.

I will now demonstrate how Sider’s commitments to Purity and Micro-Naturalism create tension for his view about reference and eligibility.

The reference relation is explanatory. That is, “one can explain certain facts by citing what words refer to.”³⁵ Moreover, because reference is explanatory, it must involve only joint-carving, i.e., natural, notions. For example, Sider (forthcoming: 37-38) says that “reference magnetism solves the problems of [bizarre interpretations] ... [as] a consequence of a more general claim that explanatory theories must be stated in joint-carving terms.” Here, Sider is making a claim that explanatory theories are joint-carving and, therefore, according to Purity, must involve only joint-carving notions.³⁶

³⁴ For a detailed discussion of naturalness, see above §3.1.

³⁵ Sider forthcoming: 32.

³⁶ As discussed above, Sider takes joint-carving notions, i.e., natural notions, to also be fundamental.

Take for example, as Sider (forthcoming: 119-123) does, the joint-carving notion of necessity. Proper application of this notion is demonstrated in (5).

(5) It's necessary that all electrons are electrons.³⁷

Here, the fundamental, i.e., natural, expression 'necessary' combines only with other fundamental expressions, i.e., 'electrons', 'also', 'all', 'are'. Conversely, sentences such as (6) are problematic applications of necessity because the fundamental expression 'necessary' is combined with a nonfundamental expression, 'cities', and is a violation of purity.

(6) It's necessary that all cities are cities.

According to Sider (forthcoming: 120), "fundamentally, there are no [city-facts]; so a fundamental notion of necessity shouldn't contain within itself the capacity to interact with such facts."

Moreover, if one accepts, as Sider claims, that reference is pure, then perfectly natural relations can have only perfectly natural relata; otherwise they would relate things that aren't perfectly natural and violate purity. Therefore, the relata of the reference relation, i.e., the referents, must be perfectly natural. Also, perfectly natural entities must be microphysical entities according to Micro-Naturalism. However, cities, Jellybeans, and Mason jars aren't microphysical, yet we seem to refer to them. This presents a problem; either (a) Sider can reject Purity, in which case natural notions are allowed to mingle with non-natural notions, or (b) Sider can reject Micro-Naturalism, in which case macrophysical entities might be natural.

This incredibly restricted notion of naturalness, i.e., one that is pure and exists only on a microphysical level, seems to cause problems for Sider's notion of eligibility and, thus, his theory of reference. Therefore, it is possible that, by appealing to an alternative notion of naturalness, we might be able to reconcile these tensions and salvage the notion of eligibility.

³⁷ Assuming that electrons exist on the fundamental level of physics.

By drawing on the work of John Hawthorne and Jonathan Schaffer, such an alternative becomes available. First, Hawthorne (2007: 434) suggests that “we should ... be willing to give relative naturalness a life of its own, one that allows properties that are of equal definitional length from the microphysical ground floor to be of radically unequal naturalness.” Rather than appealing to an austere physicalist framework that allows perfectly natural properties to exist only on the microphysical ground floor, Hawthorne (2005: 185-210) accepts an “emergentist” framework according to which naturalness “is not a matter of mere definitional distance from the microphysical ground floor.” In other words, Hawthorne would like to have a notion of naturalness that is available on various levels of nature and that does not appeal to fundamental properties. However, there is still a question of what such a notion should appeal to, if not the microphysical ground floor.

Previous notions of naturalness relied on the presence of sparse properties on the fundamental level, as is evident in Lewis’s introduction of the notion of naturalness.³⁸ However, in “Two Conceptions of Sparse Properties,” Jonathan Schaffer draws a distinction between the *fundamental* and the *scientific* conceptions of sparse properties. The former is similar to the notion of sparseness invoked by Lewis and Sider. By contrast, on the scientific conception, sparse properties are drawn from all levels of nature, not merely the most fundamental. Of these properties he says

They are those invoked in the scientific understanding of the world ... are perfectly suited to ground ... objective similarities ... are perfectly suited to carve out the causal powers of macro-properties [and] ... serve as the ontological basis for linguistic truths.³⁹

Therefore, Hawthorne’s emergentist notion of naturalness can appeal to Schaffer’s scientific conception of sparse properties. Thus, a new notion of eligibility can appeal to emergentist naturalness, which can then be applied to macrophysical entities—call this ‘Macro-Naturalism’.

³⁸ See Lewis 1986: 60

³⁹ Schaffer 2004: 92, 95, 100.

Macro-Naturalism: Natural entities can be either macrophysical or microphysical.

Thus, our new principle of reference can appeal to this new notion of eligibility—call this principle ‘Charity Plus Emergentist Eligibility’ or ‘CPEE’:

Charity Plus Emergentist Eligibility: The intended interpretation of a language L is one on which (i) referents are assigned to expressions in L in a manner that makes the sentences of a theory T true, where T is the set of sentences of L that express the beliefs of speakers of L and (ii) the relation between those speakers and the referents thus assigned is eligible, where eligibility is determined by naturalness, and naturalness need not be either fundamental or micro-physical.

This allows us to uphold Purity, reject Micro-Naturalism, and, thus, refer to macrophysical entities such as Jellybeans and Mason jars. Moreover, Macro-Naturalism is an intuitive and plausible alternative to Micro-Naturalism because (a) it seems that we do refer to macrophysical entities when we use terms like ‘Jellybean’, and (b), once we accept Macro-Naturalism, the notion of Purity and the theory that reference is a natural relation are cotenable.

I will now readdress the counterexamples provided in §4 with the new principle of reference, CPEE.

5.2 Revisiting Stranger than Fiction World, Hallucination World, and Star World

Although Macro-naturalism allows us to uphold both the principle of Purity and CPEE, it is not entirely clear that it will be able to block the bizarre interpretations presented in Stranger than Fiction World and Hallucination World Counterexamples. Unfortunately, the relations that are most eligible according to CPSE seem to *also* be most eligible according to CPEE, because there are still no alternatives. Therefore, even with CPEE, the F-interpretation and the H-interpretation are the intended interpretations for the Stranger than Fiction World and the Hallucination World, respectively. However, given that Schaffer (2007) introduces the notion of a scientific property to be a way to characterize the joint-carving nature of properties on varying levels of science (including the special sciences), it is possible that we could posit entities with scientific properties that could act as

referents for the otherwise empty names. For example, in psychology we might posit *figments* of the imagination that could act as bearers of natural properties. Thus, the object that Mario refers to when he says ‘The Mason Jar’ is merely a *figment* of his imagination.

In regard to the Star World counterexample, it is possible, according to Macro-naturalism, to have natural relations towards macrophysical entities; that is, it is possible that natural and direct perceptual relations, e.g., sight, exist between a macrophysical entity such as a planet and a speaker. Therefore, we can adopt an interpretation on which when Mario utters (4) his word ‘Betelgeuse’ refers to the planet Betelgeuse, which is the intuitive interpretation. Again, whether we allow macrophysical entities such as planets to be more natural than microphysical entities such as the photons of light traveling from planets depends on the way in which we characterize the scientific properties proposed by Schaffer (2007). For example, in astronomy we might posit planets as bearers of natural properties in order to better our “scientific understanding of the world;”⁴⁰ thus making I-interpretation more eligible than S-interpretation.

6. Conclusion

Throughout this paper, I have explored the theories presented by Lewis and Sider to find objective features that can act as semantic glue and can tie our words to the objects we take ourselves to be referring to. As demonstrated by the various principles and counterexamples to those principles, both Lewis and Sider propose a principle of reference with a notion of eligibility that rests on the assumption that natural properties must be microphysical. As shown, the alternative of Macro-naturalism and CPEE allows for an intuitively stronger principle of reference, yet it greatly depends on the way in which we formulate scientific properties. That is, CPEE depends on which properties we deem necessary for expanding our “scientific understanding of the world.”

⁴⁰ Schaffer 2004: 92

As a final note, if we are unwilling to allow for any macrophysical, scientific properties, e.g., the *figments* presented in §5.2, it is difficult for me to see how, regardless of the way in which the principle is formulated, eligibility is supposed to “save” reference in scenarios such as Stranger than Fiction World, Hallucination World, and Star World. That is, I’m not sure how an objective feature found in the world, e.g., naturalness, is going to be able to provide the means by which we tie down reference between a referrer and a referent when the intended interpretation is one on which there is no referent.

Works Cited

- Conan Doyle, Arthur. 1905. “The Adventure of Black Peter.” Reprinted in *The Return of Sherlock Holmes*. United Kingdom: George Newnes. 558-572.
- Hawthorne, John. 2005. “Epistemicism and Semantic Plasticity.” *Oxford Studies in Metaphysics* 2: 289-322. Reprinted in Hawthorne 2006: 185-210.
- . 2006. *Metaphysical Essays*. Oxford: Clarendon.
- . 2007. “Craziness and Metasemantics.” *Philosophical Review* 116.3 (July): 427-440.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, David. 1978. “Truth In Fiction.” *American Philosophical Quarterly* 15.1 (Jan.): 37-46. Reprinted in Lewis 1983b: 261-280.
- . 1983a. “New Work for a Theory of Universals.” *Australasian Journal of Philosophy* 61.4 (Dec.): 343-377. Reprinted in Lewis 1999: 8-55.
- . 1983b. *Philosophical Papers*, volume I. Oxford University Press, Oxford, New York.
- . 1984. “Putnam’s Paradox.” *Australasian Journal of Philosophy* 62.3 (Sept.): 221-236. Reprinted in Lewis 1999: 56-77.
- . 1986. *On the Plurality of Worlds*. Malden, MA: Blackwell.
- . 1999. *Papers in Metaphysics and Epistemology*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- . 2009. “Ramseyan Humility.” In David Braddon-Mitchell and Robert Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. Cambridge, MA: MIT Press. 203-222.
- Putnam, Hilary. 1977. “Realism and Reason.” *Proceedings of the American Philosophical Association* 50: 483-498. Reprinted in Putnam 1978: 123-140.
- . 1978. *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
- Sider, Theodore. forthcoming. *Writing the Book of the World*. Oxford: Oxford University Press.
- Schaffer, Jonathan. 2004. “Two Conceptions of Sparse Properties.” *Pacific Philosophical Quarterly* 84.1 (March): 92-102.
- Williams, J. Robert G. 2007. “Eligibility and Inscrutability.” *Philosophical Review* 116.3 (July): 361-398.